Supporting Information for:

In silico, in vitro and *in vivo* studies indicate the potential use of bolaamphiphiles for therapeutic siRNAs delivery

Taejin Kim, Kirill Afonin,Mathias Viard, Alexey Y. Koyfman, Selene Sparks,

Eliahu Heldman, Sarina Grinberg, Charles Linder, Robert P. Blumenthal,

Bruce A. Shapiro

RNA and DNA sequences used in this project

RNA sequences

siRNA duplexused for all experiments:
sense
5' - pACCCUGAAGUUCAUCUGCACCACCG
antisense
5' - CGGUGGUGCAGAUGAACUUCAGGGUCA

*5' sides of **sense** strand is phosphorylated.*

12-mer RNA duplex used for MD simulations:
5'-UCUUCCUGCCAUUU
5'-AUGGCAGGAAGA

Fluorescently labeled DNA sequencesused for*in vitro* FRET studies

DNA antisense_Alexa488
5' – CGGTGGTGCAGATGAACTTCAGGGTCAtt/3AlexF488N/
DNA sense_Iowa Black FQ
5' -/5IAbFQ/ ACCCTGAAGTTCATCTGCACCACCG
DNA sense_Alexa546
5' -/5Alex546/ACCCTGAAGTTCATCTGCACCACCG

Fluorescently labeled RNA sequence used for *in vivo* experiments

RNA sense_IRDye700
5`-/5IRD700/ACCCUGAAGUUCAUCUGCACCACCG

Table S1 Hydrogen bond interactions between the siRNAs and bolaamphiphiles (bolas).

	GLH-19/siRNA	**GLH-20/siRNA**
Hydroxyls(bolas)/ phosphate(siRNAs)	O1P(**rA**7)...H-O10(**b**62) [97.26]	O1P(**rG**6)...H-O5(**b**67) [99.73] O1P(**rC**4)...H-O5(**b**53) [96.41] O2P(**rC**2)...H-O5(**b**60) [65.84] O1P(**rG**28)...H-O10(**b**70) [15.70]
Bases (siRNA)/bolas	O14(**b**72)...H22-N2(**rG**33) [95.79] O3(**b**72)...H22-N2(**rG**17) [93.40] O10(**b**67)...H41-N4(**rC**26) [59.65] O12(**b**70)...H41-N4(**rC**26) [34.41] O10(**b**67)...H42-N4(**rC**26) [10.79]	O12(**b**64)...H21-N2(**rG**47) [81.55] O10(**b**65)...H22-N2(**rG**48) [48.47] O14(**b**72)...H22-N2(**rG**17) [47.74] O12(**b**65)...H22-N2(**rG**48) [19.29] O12(**b**65)...H62-N6(**rA**52) [52.82]
Sugar-edge hydroxyl (siRNAs)/bolas	O5(**b**60)...HO'2-O2'(**rU**5) [14.71] O7(**b**60)...HO'2-O2'(**rG**6) [10.80]	O5(**b**68)...HO'2-O2'(**rA**37) [38.88] O10(**b**72)...HO'2-O2'(**rG**36) [12.53]

Bolaamphiphiles and si**R**NAs are specified in parentheses by **b** and **r** respectively. In the case of siRNAs, bases (A,C,G and U) are specified. Positions of oxygen in bolas are explained in supporting Figure S2. Numbers in brackets indicate hydrogen bond occupancy for the 30-50ns range of the simulation.

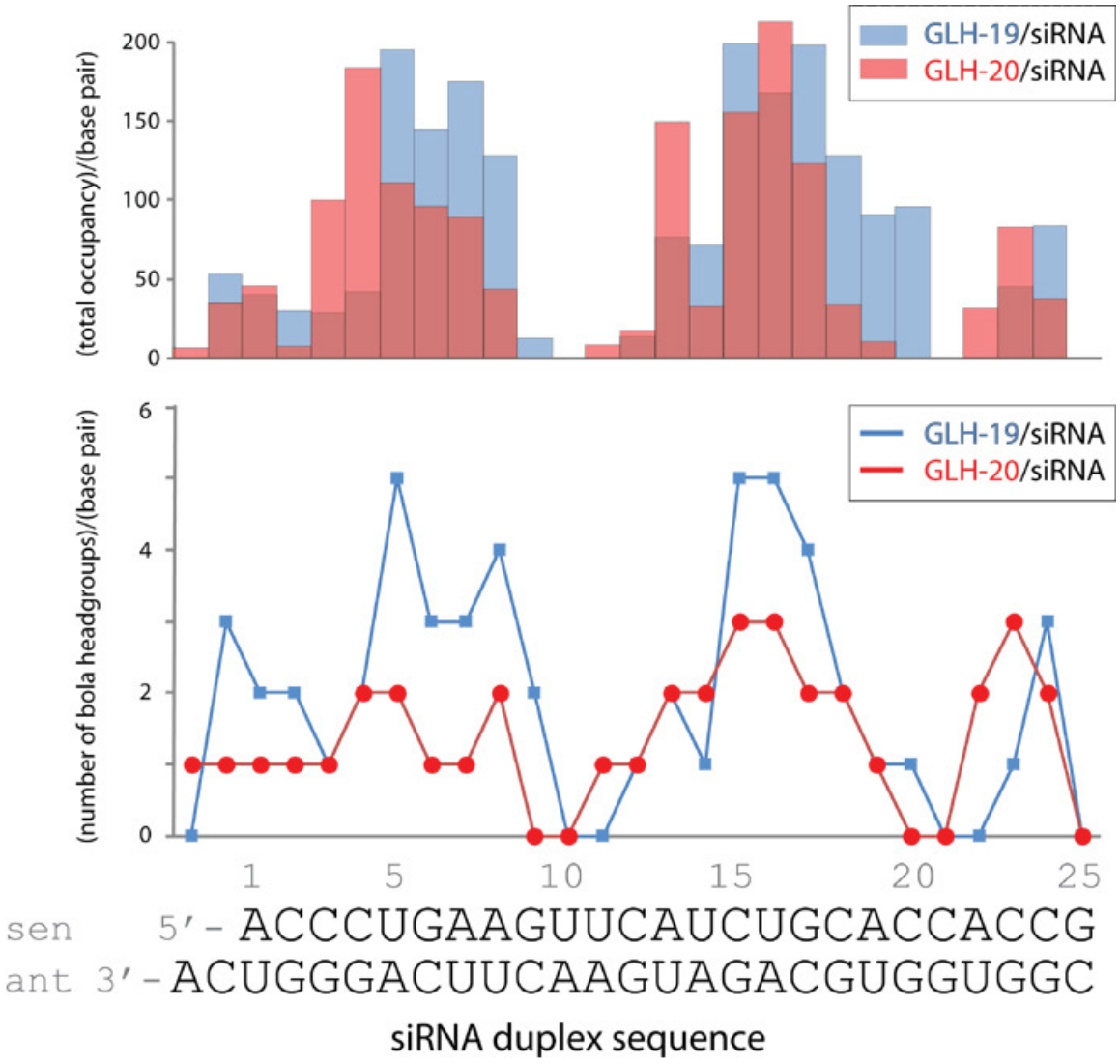

Figure S1.(**Upper panel**)Total occupancy of bola head groups near the siRNA phosphate groups in terms of base pairsfor the 30-50 ns range of the simulation. Graph indicates the summation of occupancies for each of the two phosphate groups in a base pair. (**Lower panel**)The number of bola head groups that are associated with the siRNA phosphate groups for the30-50 ns range of the simulation. Numbers above the sense strand indicate the order of base pairs from the 5' end to the 3' end. Base pairs 1-8, 12-19 and 22-25 are associated with the bola micelle surface.

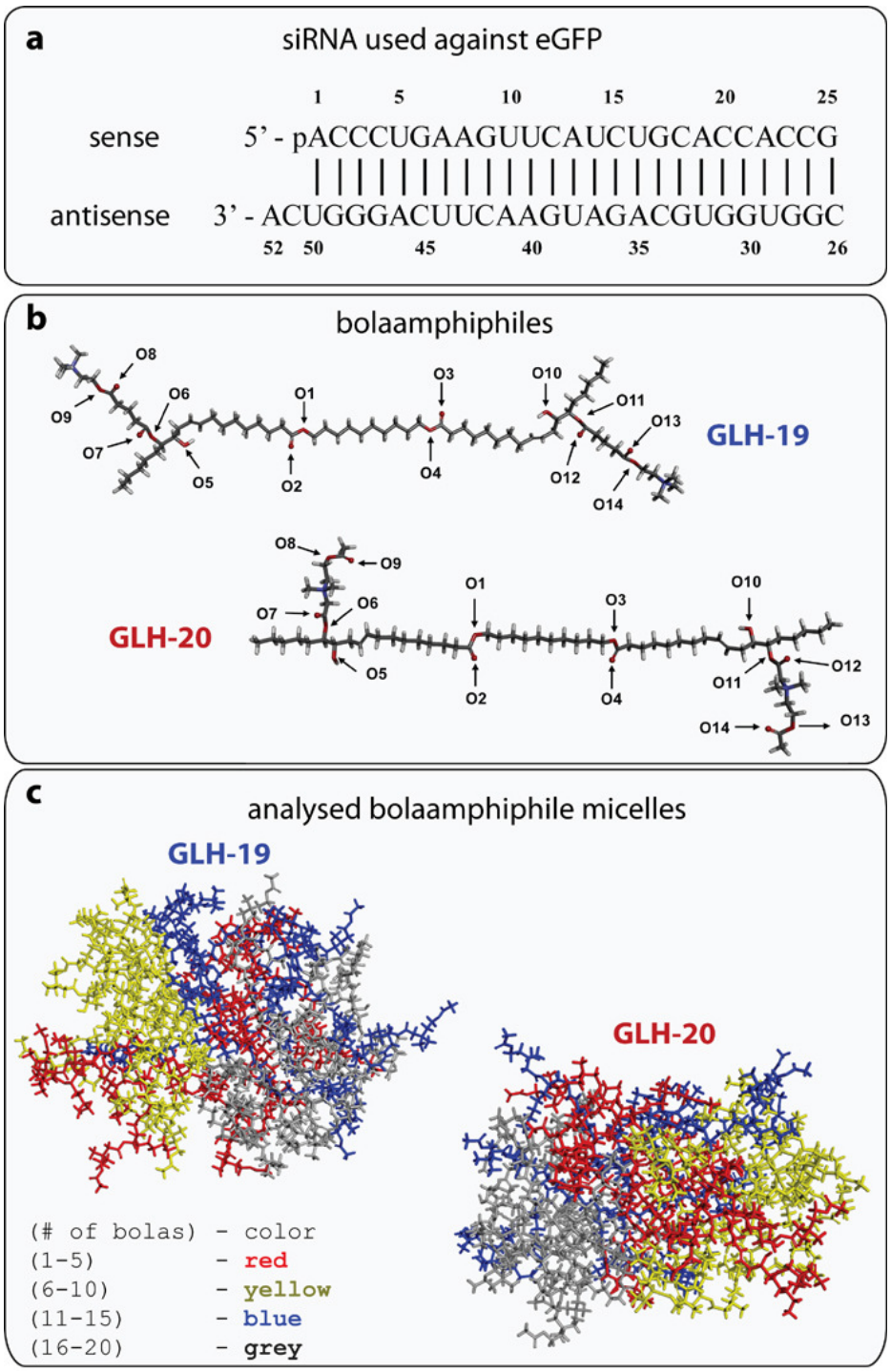

Figure S2. Structures and nomenclature of siRNA duplex and bolas used. 2D structure of siRNA duplex with corresponding numeration of nucleotides **(a)**. Bolas GLH-19 and GLH-20 with corresponding numeration of potential hydrogen bond forming atoms of oxygen **(b)**.Micelles formed by 20 bolas GLH-19 and GLH-20 with color-coded groups of bolas (1-20) used to describe the hydrogen bond formations in supporting Table S1 **(c)**.

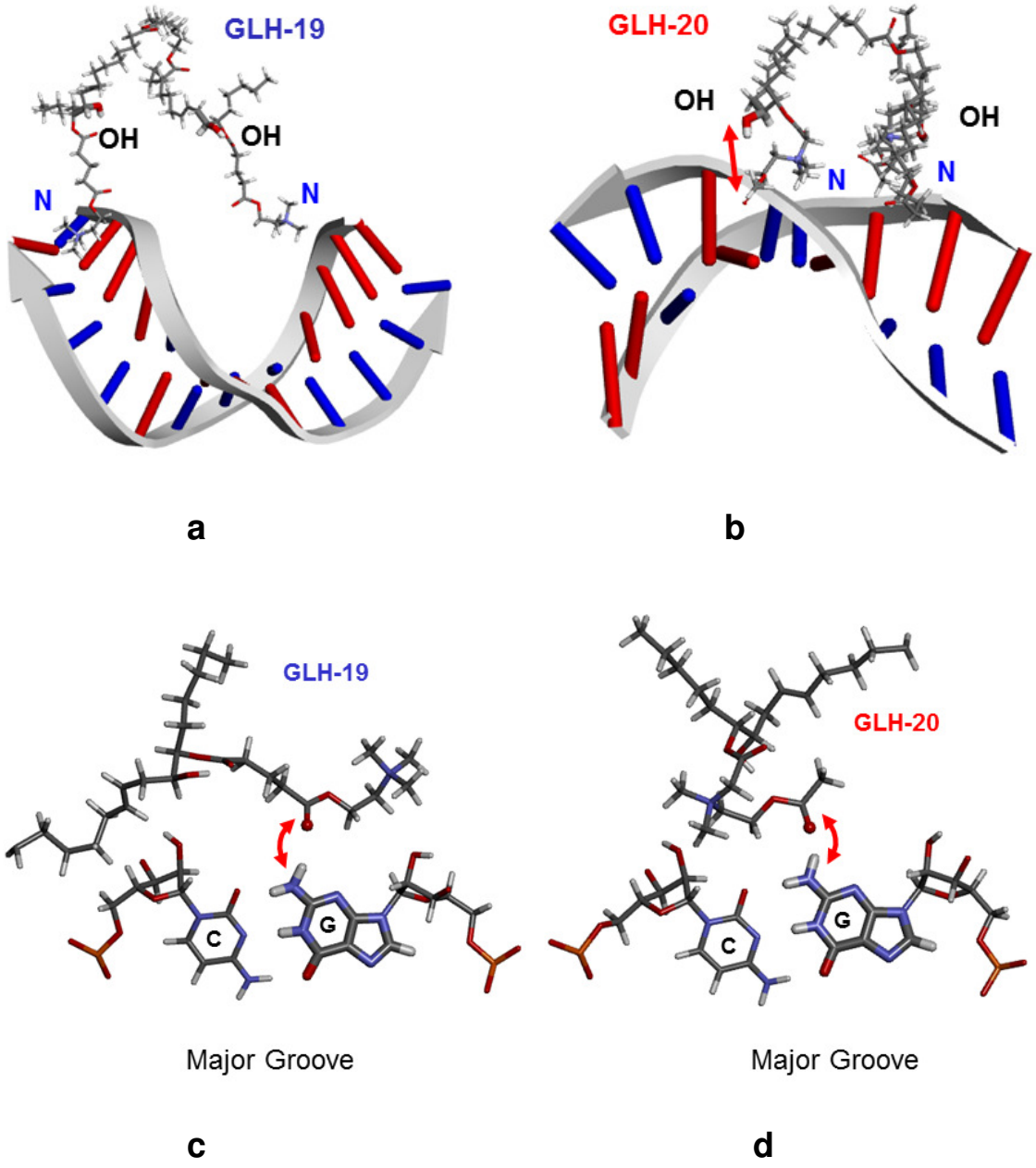

Figure S3.Examples of typical hydrogen bond interactions between hydroxyls of one bola(other bola molecules in each micelle are not shown)andthe phosphates ofthe siRNA backbonefor GLH-19 **(a)**and GLH-20 **(b)**.Specific hydrogen bond interactions between guanosine (G) and GLH-19 **(c)** and GLH-20 **(d)**.

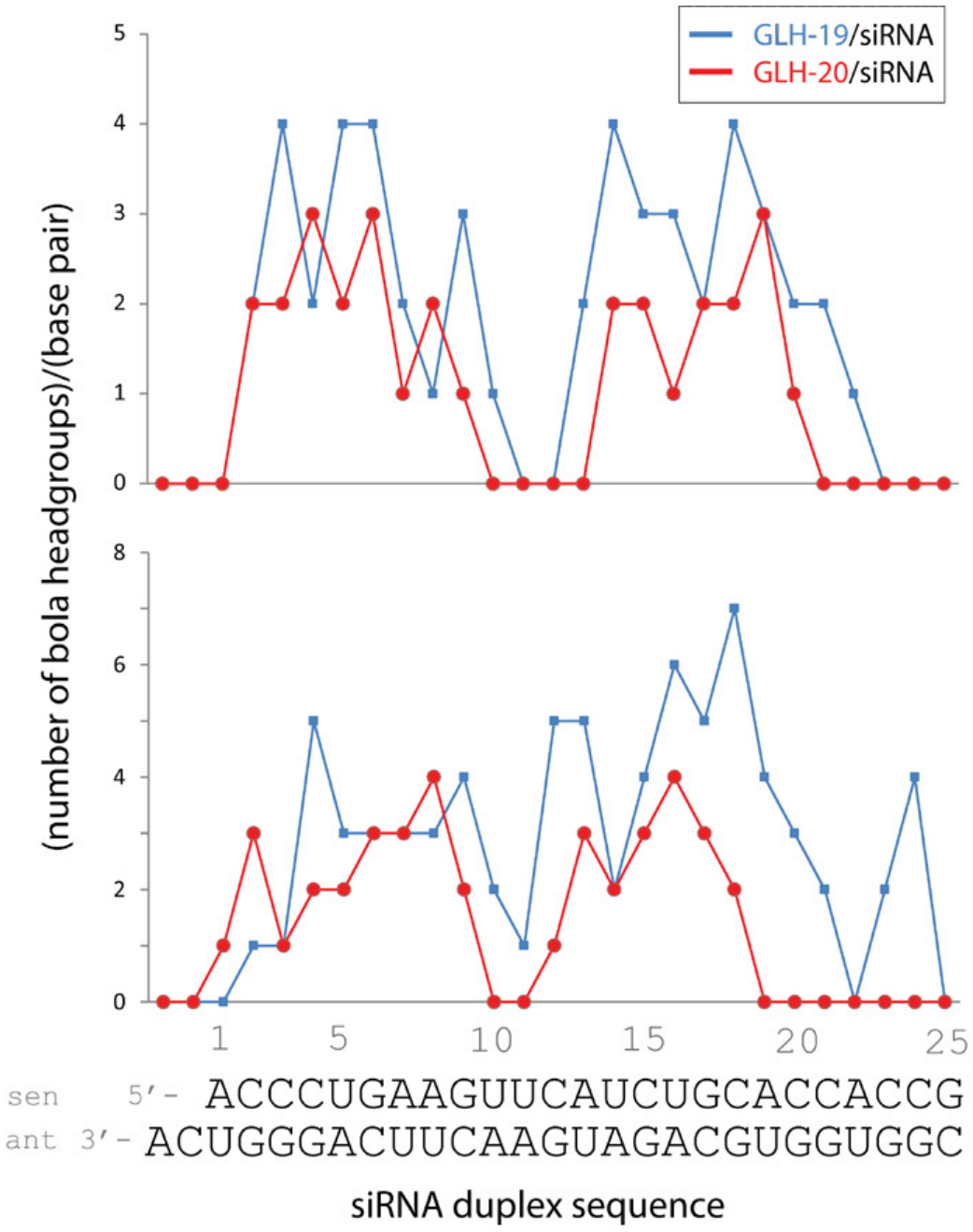

Figure S4. The number of bola head groups that are associated with the siRNA phosphate groups in the additional MD simulations, which used different velocities (**upper panel**)and different micelle surfaces (**lower panel**). Numbers above the sense strand indicate the order of base pairs from the 5' end to the 3' end.

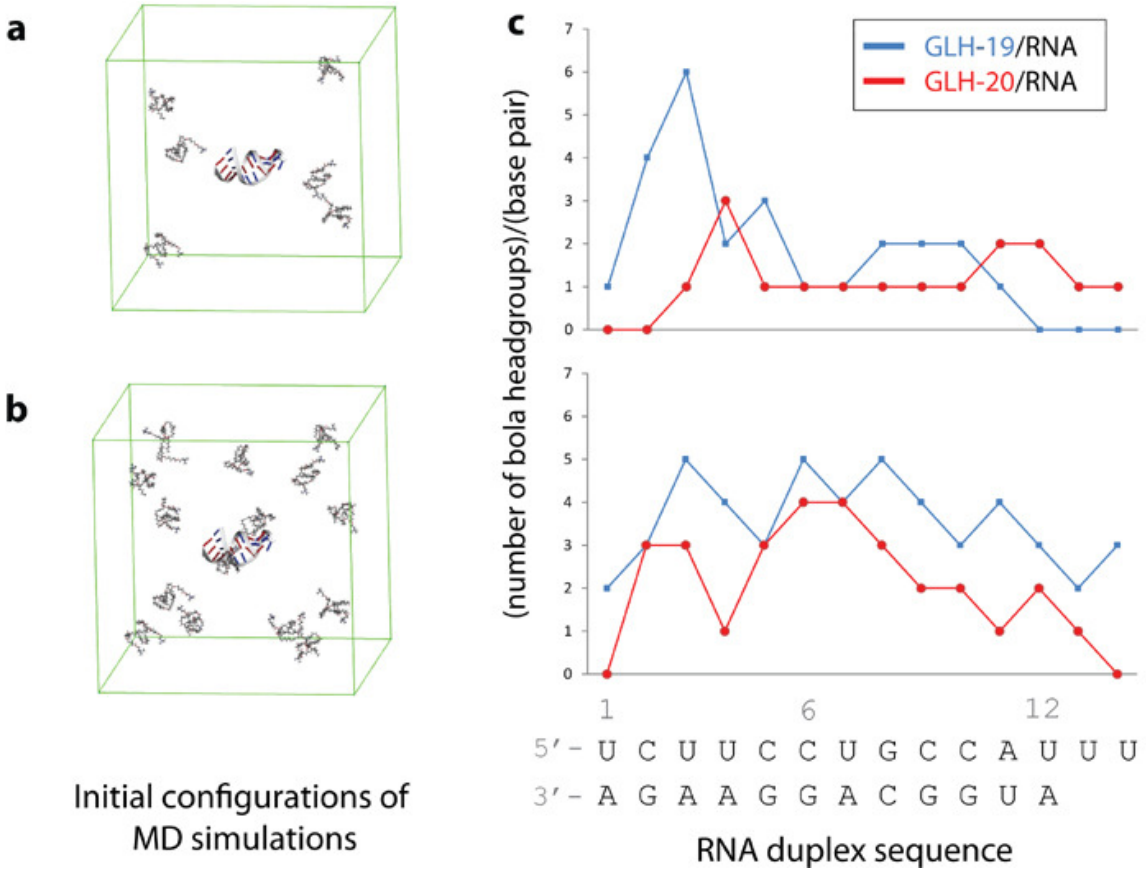

Figure S5. MD simulations of individual bolas interaction with RNA duplex.The initial configurations of MD simulations with(**a**)6, and(**b**) 15 individual GLH-19 bolas were placed around a 12-mer RNA. Green box is the size of water box (initial configurations of GLH-20 bolas are omitted). (**c**)The number of bola head groups associated with phosphate groups in the 12-mer RNA duplex for the 20-30 ns range of eachMD simulation.

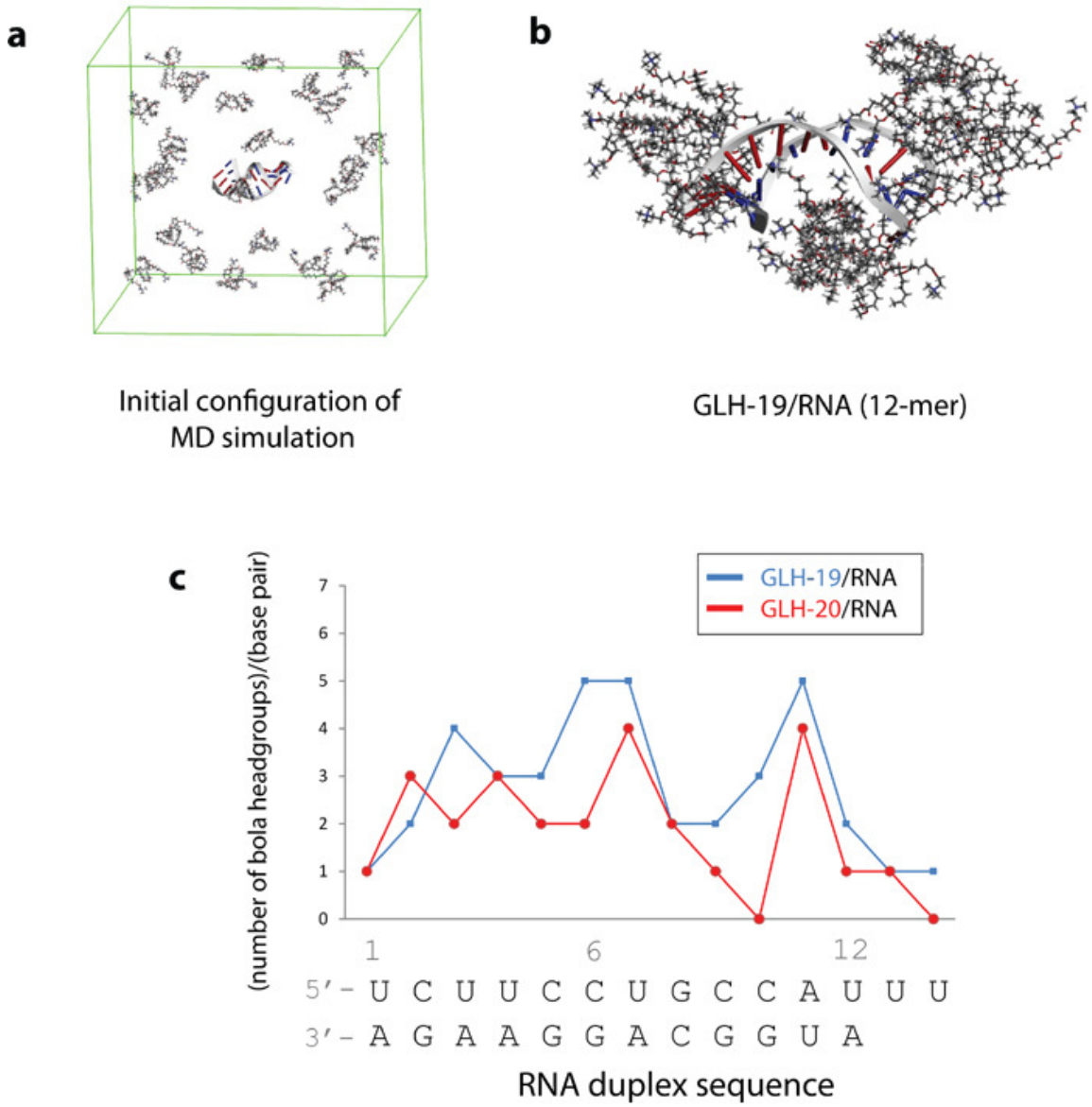

Figure S6. MD simulations of individual bolas interaction with RNA duplex.(**a**)The initial configurations of MD simulations which 24 individual GLH-19 bolas were placed around a 12-mer RNA. Green box is the size of water box (initial configurations of GLH-20 bolas are omitted). (**b**) The snapshot of GLH-19/RNA (12-mer) complex at time = 30ns. Note the formation of 3 small micelles.(**c**) The number of bola head groups associated with phosphate groups in the 12-mer RNA duplex for the 20-30 ns range of eachMD simulation.

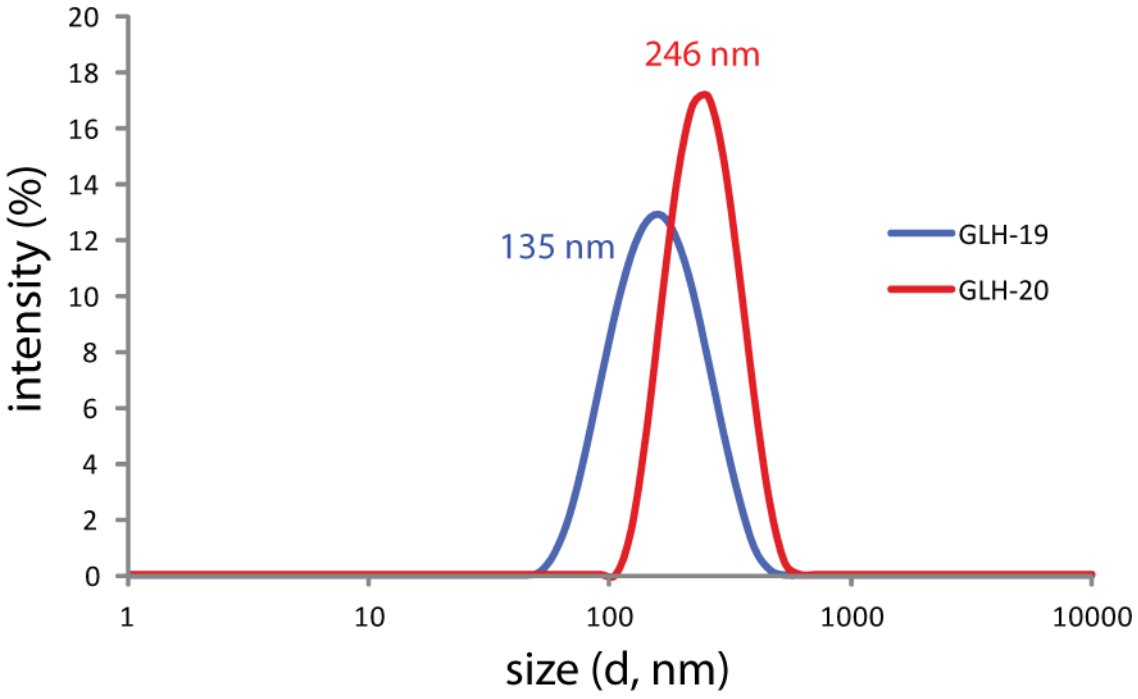

Figure S7. Size histograms from dynamic light scattering (DLS) experiments indicate average diameters for GLH-19 (blue) and GLH-20 (red) micelles to be 135 nm and 246 nm respectively.

SUPPORTING MATERIALS AND METHODS

Molecular dynamics simulations

The initial structures of the siRNA and two bolas, GLH-19 and GLH-20 were prepared with Accelrys Discovery Studio® (see Fig. 1). Electrostatic surface potential maps of GLH-19 and GLH-20 in Fig. 1 were calculated using Delphi program [1,2]. Atom and bond types for the bolas were assigned by the Amber antechamber program. The electrostatic potentials (ESP) of the bolas were obtained by performing ab initio calculations using the GAUSSIAN suite of programs. Restrained electrostatic potential (RESP) charge fitting was used to fit the ESP and the resultant partial charges were used for the bolas.

All MD simulations in our study were performed with an explicit solvent. A TIP3P water box was added so that there was a 15 Å distance from each side of the solute. MD simulations were performed using the Amber 10 molecular dynamics (MD) simulation package. The general AMBER force field (GAFF) was used for both GLH-19 and GLH-20, and the ff99bsc0 force field[3]was used for the siRNA duplex. For the ions, the following parameters were used: Na+ radius 1.868 Å and well depth 0.00277 kcal/mol; and Cl- radius 1.948 Å and well depth 0.265 kcal/mol. Since both GLH-19 and GLH-20 micelles have 40 positively charged head groups each, while siRNA has 50 negatively charged phosphate groups, GLH-19/siRNA and GLH-20/siRNA systems were neutralized by adding 10 Na+ ions to each system. In addition, we added 0.15 M of NaCl to study the interactions between the RNA and the bolas under physiological condition. An additional two initial configurations were prepared by using different initial velocities and different micelle surfaces by rotating the micelles. These configurations were tested

for 40 ns long MD simulations. We also examined the interaction between a 12-mer RNA duplex in which one strand has a two base overhang at the 3′ end (see **RNA sequence**) and individual bolas. For these MD simulations, 6, 15 and 24 individual bolas were placed around the 12-mer RNA duplex (see Fig. S5 and S6). These configurations were used for 30 ns long MD simulations.

The particle mesh Ewald summation (PME)[4]was used to calculate the electrostatic interactions. Minimization of the system (containing previously equilibrated siRNA, bola micelle, water and ions) was done using harmonic constraints on the RNA and bolas. After minimization, the system was heated to 300K while constraining the solute with a 200 kcal/(molÅ) harmonic constraint. Constraints were slowly released as the system equilibrated. In order to remove the fastest hydrogen vibrations and to allow longer simulation time steps, SHAKEwas applied to all hydrogen atoms. A constant temperature of 300K was maintained using a weak-coupling algorithm while the pressure was maintained at 1.0 Pa. Production simulations was performed for 50 ns with a 2fs timestep.

Experimental methods

Synthesis of bolaamphiphiles. The bolas, GLH19 and GLH20, were synthesized in Ben Gurion University. Briefly, the carboxylic group of methyl vernolate or vernolic acidwas interacted with aliphatic diols to obtain bisvernolesters. Then the epoxy group of the vernolate moiety, located on C12 and C13 of the aliphatic chain of vernolic acid, was used to introduce both the ACh head group on the two vicinal carbons obtained after the opening of the oxirane ring. For GLH-20, the ACh head group was attached to the vernolate skeleton through the nitrogen atom of the choline moiety (Fig. 1b). Thebolawas

prepared in a two-stage synthesis: First, opening of the epoxy ring with a haloacetic acid and, second, quaternization with the *N,N*-dimethylamino ethyl acetate. For the GLH-19 that contains an ACh head group attached to the vernolate skeleton through the acetyl group (Fig. 1a), the bola was prepared in a three-stage synthesis, including opening of the epoxy ring with glutaric acid, then esterification of the free carboxylic group with *N,N*-dimethyl amino ethanol and the final product was obtained by quaternization of the head group, using methyl iodide followed by exchange of the iodide ion by chloride using an ion exchange resin.

Duplex assemblies, bolaamphiphiles/siRNA complex formations and qualitative agarose gel electrophoresis. Single stranded RNAs and fluorescently labeled DNA strands used for stability assays and transfection efficiency were purchased from Integrated DNA Technologies, Inc. In the case of the two fluorescently labeled DNA strands, additional linkers of two nucleotides (AA) were placed before the fluorescent tags. All sequences used in this work are listed above.

The oligos (RNAs and DNAs) units at concentrations specified in the text were mixed in double-deionized water and incubated in a heat block at 95° C for two minutes. Then, the samples were removed and placed at room teperature for a period of 20 minutes. Hybridization buffer (89 mM Tris, 80 mM Boric Acid (pH 8.3), 10 mM magnesium acetate) was added to the mixtures either prior to heating, or after the step at 95° C. Prepared siRNA (or DNA) duplexes were mixed with re-suspended in water bolas (GLH-19 or GLH-20) at final concentrations specified in text, followed by 30 min incubation at room temperature. Prepared complexes were analyzed at room temperature

on 2% agarose gels in the presence of 89 mM Tris-borate, pH 8.3. A Hitachi FMBIO II Multi-View Imager was used to visualize ethidium bromide stained oligos.

Dynamic Light Scattering (DLS) experiments.For DLS experiments, 400 µl of sample solutions containing pre-formed RNA/bola complexes were used. The samples were measured at 25°C witha Zetasizer nano (Malvern Instruments Ltd) equipped with a 633 nm laser. The dynamic light scattering was analyzed by cumulant analysis. This corresponds to a polynomial fit of the log of the G1 correlation function: $\ln[G1]=a+bt+ct^2+dt^3+...$, where b is the second order cumulant, or the z-average diffusion coefficient. This value is proportional to the size of the particle. $2c/b^2$ is the polydispersity index (PDI). It can be assumed that the population is narrow and unimodal for samples with PDI values below ~ 0.2. Z average and PDI are accurate. For values of PDI between 0.2 and 0.5, samples have a distribution that is broader and more polydisperse. The Z average and PDI can be used for comparison purposes. Values of PDI over 0.5 indicate heterogeneous distribution of particles.

Stability assays with RQ1 DNase fluorescent studies.To study the stability of pre-formed duplex/bola complexes in the presence of nucleases, DNA duplexes containing one 3′ antisense strand modified with Alexa488 and a 5′ sense strand modified with IowaBlack FQ (see supporting information) were pre-incubated with bolas (GLH19 or GLH20) and the dequenching of Alexa-488 upon digestion with RQ1 DNase (Promega) was followed by FRET experiments. For all the experiments, the excitation wavelength was set at 460 nm and the excitation and emission slit widths were set at 2 nm. Upon excitation at 460 nm, the emission at 520 nm was recorded every 30 seconds to follow the digestion of the DNA duplex associated with bola. As the control, quenched

DNA duplexes were used alone. RQ1 RNase free DNase was used according to the manufacturer's protocol.

Cryogenic Electron Microscopy (cryo-EM) experiments. Quantifoil Copper 200 mesh R 3.5/1 grids were washed overnight with acetone. To prepare a frozen, hydrated grid, 2.5 μL of sample was applied to the grid, blotted, and plunged into liquid ethane using Vitrobot III (FEI, Hillsboro, OR). Images were collected at liquid nitrogen temperature (~100 K) on a JEM-2010F (JEOL Inc., Tokyo, Japan) transmission electron cryo-microscope equipped with a field emission gun (FEG). JEM-2010F was operating at 200 kV and was equipped with a Gatan cryo-holder (model 626) (Gatan Inc., Pleasanton, CA). Images were recorded on DE-12, a 12.6 megapixel (3072 X 4096) Direct Detection Device (DDD) sensor (Direct Electron LP, San Diego, CA). Samples were imaged at 13900X effective magnification targeted at 3–4 μm underfocus. We used a total specimen exposure for each image of 30 $e^-/Å^2$sec.

Transfection of human breast cancer cells withbolaamphiphile/siRNA complexes.For assaying the delivery of functional siRNAs, human breast cancer cell line *MDA-MB-231* (with or without eGFP)was grown in D-MEM media (Gibco BRL) supplemented with 10% FBS and penicillin-streptomycin in a 5% CO_2 incubator. All transfections in this project were performed using bolas (GLH19 or GLH20). The concentration of siRNA (500 nM) to bolas(500 μg/ml)was 50X. siRNAs/bolas solutionswere pre-incubated at 30^oC for 30 min. Prior to each transfection, the cell media was swapped with OPTI-MEM and prepared 10X or 50X siRNA/bola complexes werediluted to the final concentration of 1X. The cells were incubated for 4 hours followed by the media exchange (D-MEM, 10%FCS, 1% pen-strep).

Fluorescent LightMicroscopy.To assess the silencing efficiency, cells were imaged 72 hours after the transfection with a Nikon 200 TE inverted microscope (Melville, NJ). We used a PanFluor 20x, ELWD, NA=0.45 objective and a Nikon B-2E/C, 465–495/505/515–555 cube for eGFP imaging (Chroma Technology Corp., Rockingham, VT). MetaMorph software (Universal Imaging Co., Downingtown, PA) was used to operate the microscope.

Flow cytometry experiments.For statistical analysis of the flow cytometry experiments, the *MDA-MB-231* cells (with or without eGFP) grown in 12-well plates ($10x10^4$ cells per well) were lifted with cell dissociation buffer, washed twice with PBS and the level of expression of eGFP was determined by fluorescence-activated cell sorting (FACS) analysis on a FACScalibur flow cytometer (BD Bioscience). At least 30,000 events were collected and analyzed using the Cell quest software.

In vivo experiments.Animal studies were performed according to the Frederick National Laboratory for Cancer Research (Frederick, MD) Animal Care and Use Committee guidelines. Imaging studies were performed on MDA-MB-231 tumor bearing athymic nudes (Charles River Laboratories, Frederick, MD). For bio-distribution experiments: after sufficient growth of injected MDA-MB-231 tumors (~ 2 weeks), the mouse was injected in the tail vein with siRNA_IRDye700 associated with GLH-19. Fluorescence imaging (Maestro GNIR-FLEX, Cambridge Research & Instrumentation, Inc. Woburn, MA) was performed at baseline (pre-injection for determining autofluorescence) and 1 hr post injection while the animal was anesthetized (1-2% isoflurane in O_2 at 1 L/min flow). The animal's internal temperature was maintained prior, during the scan (heated imaging table), and post imaging while the animal

recovered from anesthesia. Image analysis (image library for autofluorescence and contrast agent) was performed according to manufacturer's protocol (Maestro software 2.10.0, CRi, Woburn, MA). Due to the IR wavelength parameters of the contrast agent, image acquisition utilized an excitation filter (590±15 nm), emission filter (645 nm longpass) and a multispectral acquisition of 650-850 nm with 10 nm steps. The mouse was euthanized (CO_2 asphyxiation as per ACUC guideline) and the relative uptakes of the fluorescent tags by the internal organs were estimated by microscopy.

References

1. Rocchia, W, Alexov, E, and Honig, B (2001). Extending the Applicability of the Nonlinear Poisson−Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions.*J. Phys. Chem. B* **105**: 6507–6514
2. Rocchia, W, Sridharan, S, Nicholls, A, Alexov, E, Chiabrera, A, and Honig, B (2002). Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *J. Comput. Chem.* **23**: 128–137
3. Cheatham, TE 3rd, Cieplak, P,and Kollman, PA (1999).A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *Journal of biomolecular structure & dynamics***16**: 845-862
4. Essmann, U et al(1995).A smooth particle mesh Ewald method. J. Chem. Phys. **103**: 8577-8593.